Weed Science, 50:215–226. 2002

Contemporary methods to investigate seed and bud dormancy

Wun S. Chao
USDA—Agricultural Research Service, Plant Science Research, Biosciences Research Laboratory, Fargo, ND 58105-5674; chaow@fargo.ars.usda.gov

Dormancy is a state of shifted physiological activities with cessation of growth. It occurs in seeds and vegetative propagules and enables plants to survive in adverse growing conditions. Traditional studies on dormancy-related problems have mostly focused on hormone changes along with environmental factors that have achieved great insight on these processes at the physiological level. The molecular nature and cellular basis of signals that carry out the processes of dormancy or dormancy breaking are largely unknown. Recent advances in plant genetics and genomics have provided assorted ways to investigate questions concerning dormancy. Various approaches such as developing genetic maps with DNA-based markers, e.g., amplified fragment length polymorphism (AFLP) and restriction fragment length polymorphism (RFLP), analyzing mutant lines, conducting quantitative trait loci (QTL) analysis, two-dimensional polyacrylamide gel electrophoresis, differential display, microarray, have been performed to resolve different issues related to dormancy. The phenotypic variation in dormant seeds or buds is continuous instead of discrete, and thus QTL analysis is desirable to identify the association between genetically determined phenotypes and specific genetic markers (RFLPs). Some aspects of QTL will be introduced. DNA microarray is a recently developed technology that is used to detect and quantitate large numbers of differences in gene expression simultaneously. We have used the DNA microarray technology to study underground bud dormancy and growth in leafy spurge (*Euphorbia esula* L.). The principle and versatility of DNA microarray will be introduced, and the strategy for applying this technology will be discussed.

Nomenclature: Leafy spurge, *Euphorbia esula* L. EPHES.

Key words: Genomics, dormancy, microarray, quantitative trait loci, virus-induced gene silencing.

Genomics (the efforts to map and sequence all genes and to unravel their function) has enormous impact in crop improvement. With the development of many new and better technologies, genomic research has generated an immense amount of information. Therefore, it becomes the responsibility of scientists to apply it creatively. The basic avenue to achieve crop improvement begins with gene identification, which is to locate the gene that is responsible for a particular trait along the chromosome and to examine its potential role. This approach relies on the altered phenotype because of mutation at a specific site in a chromosome, and thus is referred to as mutation-based gene identification. Once this information is acquired, we can then design experiments to clone the gene and make use of it. However, such information is not always at hand. For example, we may want to search for genes regulating cellular activities during seed development, whereas no mutant line is available for mapping purposes. Thus, an alternative approach is needed, which is to investigate changes in gene expression in response to internal or external signals at the cellular or tissue level. Because gene identification for this approach relies almost entirely on the levels of transcription or translation associated with developmental or environmental stimuli, it is referred to as expression-based gene identification. These two approaches are remarkably different, and thus many contemporary techniques are only useful for either a mutation- or expression-based approach in searching for and cloning of genes, but seldom for both (Table 1).

In this paper, some of the most popular and effective molecular tools used in genetic and genomic research will be outlined, and their application to unravel biological processes such as bud and seed dormancy will be described. Two important molecular tools, microarray and quantitative trait loci (QTL) analysis, will be discussed in some detail because they represent important progress in genomic research, yet are performed based on different research principles (expression-based vs. mutation-based). Current molecular tactics to determine the function of genes will also be presented. One recently developed method that is based on the concept of targeted gene silencing via a genetically modified virus will also be discussed in some detail. This method is important because of its potential for examining the function of a large number of genes in a relatively short period of time.

This article is not meant to provide a thorough and detailed discussion of each technique, but rather to give concise principles on how these techniques work in terms of problem solving and their advantages or disadvantages. The content is intended for readers without a strong molecular background. However, if readers are not familiar with the basic ideas of cell biology and the concepts of gene action, they are advised to read a basic biology and genetic textbook (Buchanan et al. 2000; Lewin 1997).

Contemporary Methods

Dormancy is a physiological state with cessation of active growth. Dormancy enables plants to survive in adverse

TABLE 1. Common techniques used for mutation- and expression-based gene identification and their general differences.

Gene identification	Expression based	Mutation based
Methods	Serial analysis of gene expression differential RNA display, microarray, two dimensional polyacrylamide gel electrophoresis, etc.	Developing genetic maps, quantitative trait loci analysis (QTL), gene traps, T-DNA knockout mutants, etc.
Approaches for function determination	Reverse genetics approach: loss-of-function phenotype, over-expressing phenotype	Forward genetics approach: functional complementation
Requirements	Need to have at least two different experimental conditions (i.e., induced and control) for comparative studies	Need a genetic linkage map for QTL analysis and positional cloning

growing conditions. Various research on seeds and buds (underground adventitious buds and axillary buds) have demonstrated that the mechanism of dormancy is complicated and involves many genes (Anderson et al. 2001; Foley 2001). To elucidate the molecular and cellular basis of signals that carry out the processes regulating dormancy, cloning of genes involved in dormancy processes is undoubtedly required. Table 1 presents two groups of techniques based on similarity of principles and specifies differences between the two.

Mutation-based Gene Identification

Basically, genes responsible for specific phenotypes such as early, late, or loss of germination may be investigated by identifying mutations at the DNA level. Mutation analysis is a relatively direct way of monitoring gene effects as a mutation either occurring naturally or artificially reflects the function of that gene product. Among the different methods listed in Table 1, the information obtained from genetic map and QTL analysis can be used for map-based cloning. Recent studies on dormancy indicated that seed germinability or bud arrest is not regulated by qualitative genes but rather by many QTL (Foley 2001; Frewen et al. 2000; Paterson et al. 1995). Thus, determining the sites of QTL responsible for the regulation of dormancy in the chromosome is a prerequisite for further cloning and for deciphering the mechanistic roles of these genes. The concept of QTL will be discussed later.

To date, T-DNA knockouts and gene traps have not been used for investigating dormancy-related questions, probably because these methods are not ideal for studying quantitative traits. T-DNA knockouts are developed by inserting transferred DNA or T-DNA into the plant genome. This foreign DNA disrupts the expression of the gene where it is inserted and serves as a marker for subsequent identification of the altered gene (Krysan et al. 1999). The T-DNA–transformed *Arabidopsis* lines are accessible to the public through the *Arabidopsis* Knockout Facility at the University of Wisconsin (http://www.biotech.wisc.edu/arabidopsis/). Gene traps use T-DNA or a transposable element as the insertion vehicle and are similar to T-DNA mutation in many aspects. The major difference between the two is that gene traps contain a reporter gene. Gene traps identify genes through the specific patterns of reporter gene expression with or without phenotypic alteration (Springer 2000). These two techniques are most beneficial to scientists who use *Arabidopsis* as research material because the whole genome has been sequenced (Ausubel 2000). Both T-DNA mutation and gene traps are limited to plant species that can be easily transformed and regenerated. In addition, because they involve the random insertions of DNA into the chromosome, tens of thousands of transgenic lines are needed to cover the whole genome. Thus, screening for specific traits is not easy.

Expression-based Gene Identification

An expression-based approach identifies genes solely by differentially expressed mRNA or proteins under various experimental conditions. Table 1 lists a few methods that are widely used. Serial analysis of gene expression (SAGE) allows the quantitative and simultaneous analysis of a large number of transcripts. This technology is based on three principles: (1) a short sequence tag (10 to 14 bp) is generated that contains sufficient information to uniquely identify a transcript, provided that the tag is derived from a distinctive position within that transcript; (2) many transcript tags can be concatenated or linked into a long, single molecule that can be cloned and sequenced; and (3) the relative abundance of each expressed gene can be obtained by the number of times the corresponding tag is observed (Veculescu et al. 1995). This method is very efficient for examining gene expression levels because it is able to analyze a large number of different RNA species in a short time. However, its usefulness is limited to some extent because the 10 to 14 bp tag is too short to be used for cloning purposes on a routine basis.

Differential screening of a cDNA library (a collection of cDNA clones) and differential display are excellent tools to distinguish mRNA in comparative studies. When differentially screening a cDNA library, RNA isolated from test and reference samples are used to generate radioactive cDNA probes and are hybridized independently to duplicate filters derived from the same cDNA library (Sambrook et al. 1989). Clones that hybridize to both probes indicate that the corresponding genes are expressed in both conditions, whereas clones that only hybridize to one probe indicate that the corresponding genes are only expressed in one condition. The disadvantage of this method is that rare mRNA will have very low specific probe concentrations, and thus might not detect the corresponding cDNA clones after hybridization.

Differential display is a polymerase chain reaction (PCR)-based method (Liang and Pardee 1992, 1995). The plant RNA is first synthesized to its cDNA form and then used as a template for the PCR reaction using two types of PCR primers, anchored oligo(dT) and arbitrary decamer. The amplified cDNA products are separated by polyacrylamide gel electrophoresis (PAGE), and the patterns of bands are compared between the test and reference samples. Because

differential display relies on the power of PCR amplification, it is much more sensitive than that of differential screening of a cDNA library; however, it is also prone to a high false positive rate (Li et al. 1994; Sun et al. 1994). In addition, because of the sensitivity of PCR amplification, smaller amounts of the starting material are needed. The downside of both methods is that only a limited number of clones can be studied at one time, and the work can be very laborious if the investigation requires collecting numerous samples from plants grown under various environmental conditions and in various stages of development, as do most research problems.

In contrast, microarray technology (Duggan et al. 1999; van Hal et al. 2000) has the capability to simultaneously detect and quantify a large number of differences in gene expression (global gene expression) related to pathways and mechanisms governing dormancy. In addition, this technology becomes particularly important to study some organisms with poorly defined genetics as in the case of investigating crown and root bud dormancy in leafy spurge (*Euphorbia esula* L.). The microarray technique will be discussed later.

Recent evidence indicates that mRNA-based analysis alone does not provide sufficient information about gene expression because there is a discrepancy between the relative expression levels of mRNA and their corresponding proteins (Gygi et al. 1999). To make a complete analysis of gene expression, a protein-based study in conjunction with analysis of mRNA expression levels is needed. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is commonly used for analyzing protein expression. This technology is now developed into a new arena called proteomics (study of global protein expression). Basically, proteins are separated by immobilized pH gradient (IPG) in the first dimension, and by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) in the second dimension. The unique 2D-gel spots are then characterized using mass spectrometry or Edman N-terminal sequencing (Dutt and Lee 2000). This technology has been advanced to a point that one can almost identify any protein as long as it can be observed on a gel. Most physiological responses are manifested at the level of protein activity. Thus, protein data will be an important adjunct to our endeavors to decipher gene expression at the transcriptional level.

Microarrays

Microarray technology has been developed and widely used for investigating diverse questions in plant and animal biology in the last 5 yr (van Hal et al. 2000). Microarray technology works on the principle that cognate nucleic acids hybridize with each other. This technology allows biologists to systematically evaluate the expression pattern of large subsets of genes at once in given tissues over multiple developmental stages and in response to various environmental stimuli (Duggan et al. 1999; van Hal et al. 2000). The technique was originally developed based on direct synthesis of oligonucleotides on a solid surface (Foder et al. 1991). However, the design of this type of array (DNA chip) is so expensive that it is not accessible to most labs. On the other hand, cDNA-based microarray, first developed by Schena et al. (1995), is made by arraying a large number of cDNA on a glass or membrane support and then simultaneously hybridizing with fluorescently tagged cDNA pools from test and reference samples.

Figure 1 illustrates the practical application of microarray analysis. DNA is prepared by PCR amplification of cDNA clones using 96-well plate formats. Following purification and quality control, the amplified cDNA are printed on a coated glass slide using a robotic arrayer such as a Cartesian Robot (Seiko Instruments, United States). To compare the relative abundance of mRNA between test and reference RNA samples, the two samples are labeled separately with different fluorescent tags (e.g., reference mRNA with red tag and test mRNA with green tag), mixed, and then hybridized with the DNA spots on the microarray. After hybridization, the fluorescent DNA spots are measured using two laser light sources (ScanArray 3000) for detecting red and green fluorescence. The relative abundance of each gene in these two RNA samples is reflected by the ratio of red to green fluorescence.

DNA microarray images can be analyzed by ScanAlyze2 (http://www.bio.davidson.edu/Biology/GCAT/protocols/scanalyze.html), which is accessible to academics at no charge. ScanAlyze2 quantifies the intensity of fluorescence from each spot. To analyze the coordinately regulated genes from a series of array hybridizations, CLUSTER and TREEVIEW programs can be used. These two programs were developed by Eisen et al. (1998) and are also accessible to academics at no charge. CLUSTER compares each variable gene to another variable gene from a series of array hybridizations and groups genes with similar expression patterns. TREEVIEW displays the graphic form of the CLUSTER analysis. Based on clustering expression results from yeast, mice, and human genes, two groups of scientists found that genes participating in similar processes or pathways often share similar regulatory mechanisms, resulting in comparable expression profiles (Eisen et al. 1998; Wen et al. 1998). The power of these results is that clustering information can associate poorly characterized genes with genes whose function and regulation are known and thus provide a direction for scientists to study the poorly characterized genes (Eisen et al. 1998; Wen et al. 1998). Recently, new evidence indicates that cluster analysis will help elucidate the regulatory structure of genomic networks based on motifs matched from the upstream of open reading frames within each cluster (Tavazoie et al. 1999). Many other programs for microarray gene expression analysis are also available at some cost (Stevenson 1999).

Microarray technology has been widely received and used to identify genes of specific functions (Aharoni et al. 2000), to compare transcript profiles under different environmental conditions (Reymond et al. 2000), to evaluate transcript profiles between genetically modified and control species (van Hal et al. 2000), and to characterize differentially expressed genes between tumor and normal cells (Epstein and Butow 2000). The limitation of this technology is that it requires a sufficient number of cDNA clones to be printed on the slide and expensive facilities to perform the tasks. Recently, Dr. David Horvath (personal communication), USDA-ARS, Fargo, ND, has used *Arabidopsis* microarrays to identify differentially expressed genes from leafy spurge and other plant species. The results showed that fluorescent-tagged leafy spurge cDNA hybridized to 60% of the 12,000 *Arabidopsis* cDNA clones on the microarray. These results

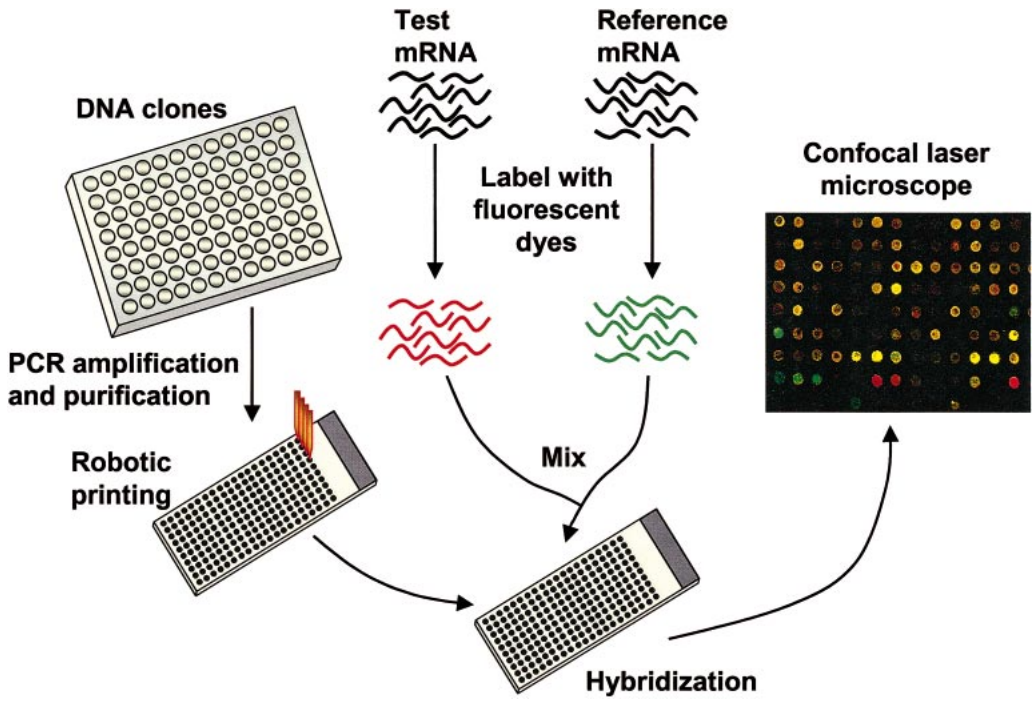

FIGURE 1. Gene expression analysis using a cDNA microarray. DNA is prepared by PCR amplification of cDNA clones using 96-well plate formats. Following purification and quality control, the amplified cDNA are printed on a coated glass slide using a robotic arrayer. To compare the relative abundance of mRNAs between test and reference RNA samples, the two samples are labeled separately with different fluorescent dyes (e.g., reference mRNA with red dye and test mRNA with green dye), mixed, and then hybridized with the DNA spots on the microarray. After hybridization, the fluorescent DNA spots are measured using two laser light sources for detecting red and green fluorescence. The relative abundance of each gene in these two RNA samples is reflected by the ratio of red to green fluorescence.

open the possibility that microarray technology can be applied to study gene expression from any plant species using ready-made slides that are printed with cDNA clones derived from convenient sources such as *Arabidopsis*.

Quantitative Trait Loci

Many important agricultural traits such as crop yield, grain weight, and plant height are quantitative traits. They are controlled by multiple genes (polygenes) and display continuous phenotypic variation for the trait (Stansfield 1991). The loci (see definition later) controlling these traits are called QTL. The primary purpose of locating QTL is to use them as selection tools for plant breeding (marker-assisted selection). Traditionally, plant breeders will cross two parents and perform selection until the lines with the best phenotypic trait are identified. These lines will be further evaluated by serial trials so that the best lines can be released to farmers as a new cultivar. This type of selection requires a large input of time, labor, money, and land. On the other hand, the selection procedure will be much more efficient if breeders have identified the QTL with nearby molecular markers. Breeders can evaluate hybrid seedlings using molecular markers that are known to be closely associated with the trait of interest and select only those having these marker alleles. In this way they do not need to wait for crops to mature before performing selection and thus save time, resources, and space. Marker-based selection also completely excludes the variation caused by environmental factors. Moreover, if the precise locations of QTL are identified it would open up the possibilities of positional gene cloning, and characterization and future use of cloned genes. The following sections introduce the principle of major gene and QTL mapping.

Genetic Terminologies

The definitions of a few frequently used genetic terms (locus, allele, molecular marker, and linkage, genetic, and physical maps) are provided below. Locus (the plural is loci) is the position on a chromosome at which the gene or one of its alleles (alternative forms of the same gene) resides. Markers (molecular and morphological markers) are used to construct a linkage, genetic, or physical map. Thus, a marker is an identifiable site on a chromosome whose inheritance can be monitored by segregation analysis. Molecular markers are much more abundant than morphological markers. In addition, molecular markers, unlike morphological markers, do not show phenotypic variations and are not affected by the physiology of the organism, and thus they are excellent tools for QTL mapping. Four types of molecular markers have been widely used to generate linkage maps, restriction fragment length polymorphisms (RFLPs), simple sequence length polymorphisms (SSLPs), random amplified polymorphic DNA (RAPDs), and amplified fragment length polymorphisms (AFLPs) (Sambrook et al. 1989; Tautz 1989; Williams et al. 1990; Zabeau and Vos 1993). The functional relationships among these markers are well explained by Brown and Kresovich (1996) and Paterson (1996a).

Mapping places molecular and phenotypic markers in order on a chromosome. These markers will then act as signposts to point to the position in relation to other important morphological trait loci for cloning or selection purposes. There are three different types of maps: linkage, genetic, and

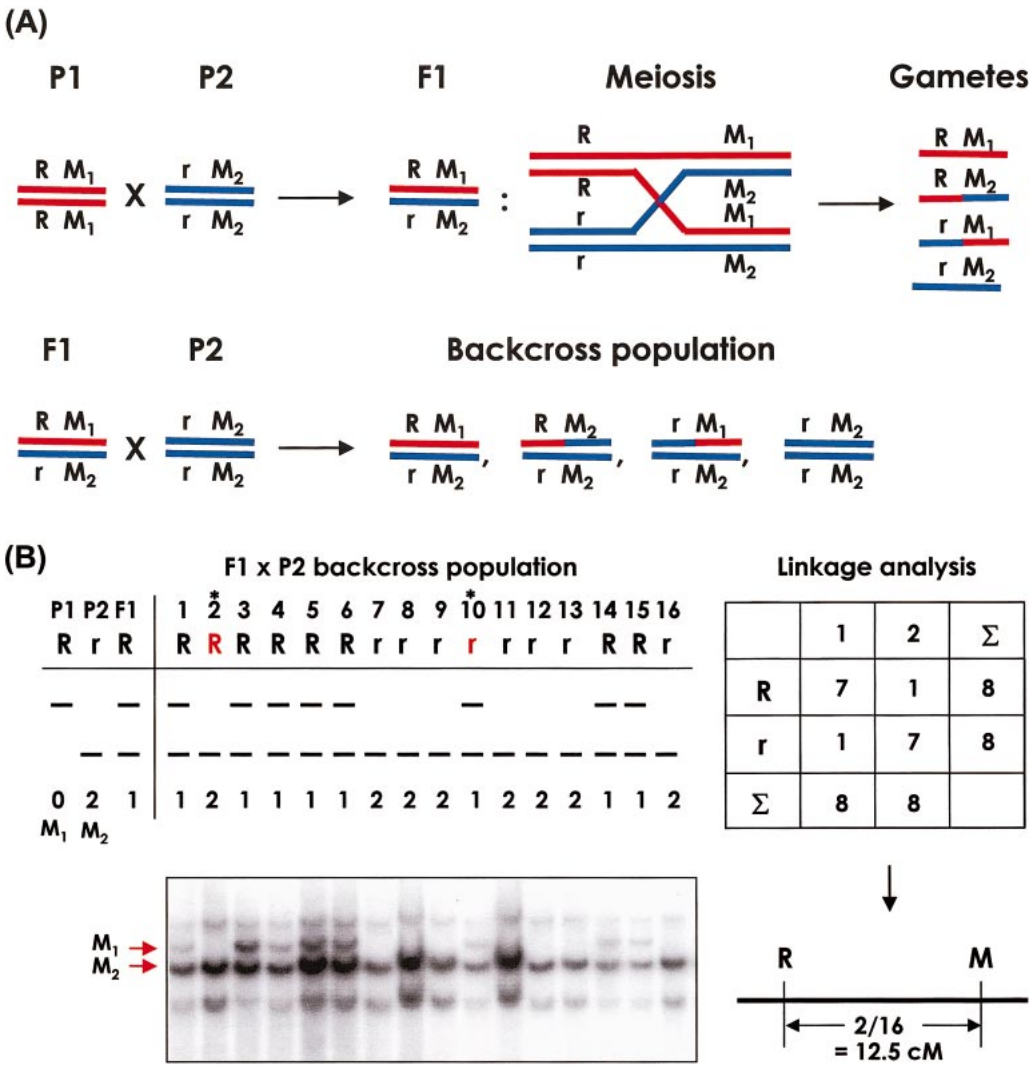

FIGURE 2. The distance between a major gene and a molecular marker (RFLP) is determined by the percentage of crossover. (A) M_1 represents marker allele 1 which segregates together with the disease-resistant (R) allele in parent 1 (P_1). M_2 represents marker allele 2, which segregates together with the disease-susceptible allele (r) in parent 2 (P_2). A cross between two homozygous parents ($RM_1/RM_1 \times rM_2/rM_2$) produces a heterozygous (RM_1/rM_2) F_1 generation. The F_1 segregates and gives rise to four different gamete types (RM_1, RM_2, rM_1, and rM_2). The genotype of these gametes can be identified by carrying out a testcross between F_1 and the disease-susceptible parent, P_2. Four classes of testcross progeny are obtained (RM_1/rM_2, RM_2/rM_2, rM_1/rM_2, and rM_2/rM_2), in which RM_2/rM_2 and rM_1/rM_2 are the recombinants. (B) Genomic DNA of 16 backcross progeny are digested with a restriction enzyme, run on an agrose gel, blotted onto a membrane, and hybridized with a probe that detects the single fragment with identical (or nearly identical) sequences. R and r represent plant phenotype. 1 and 2 represent marker phenotype. Nonrecombinant plants are R/1 or r/2. Recombinant plants are r/1 or R/2. In this exercise, two individual progeny (line 2 and line 10) have recombined the original linkage relationships of the parent. These two individuals are called crossover type, and the distance between the locus R and M is determined by the percentage of crossover. cM: centimorgan.

physical. A linkage map is a map of a chromosome, showing in linear order the relative positions of known genes (or markers) based on recombination (or crossover) frequencies. A genetic map provides a linear order and the relative distances between mutant sites on a chromosome based on various recombination frequencies. A physical map identifies the physical location of genes on a chromosome and is generated based on pulsed field gel electrophoresis, fluorescence in situ hybridization, and contig mapping. During meiosis, the paternal and maternal chromatids join and form chiasmata (the singular is chiasma), and crossover takes place through the breakage and reunion of nonsister chromatids within a chiasma. Although chiasmata may occur anywhere along a chromosome, their distribution is not uniform, resulting in a distortion in the relationship between the physical map and the genetic map (Paterson 1996a).

Determining the Distance Between a Major Gene and a Molecular Marker

Major genes are inherited in a Mendelian manner, and their allelic forms give rise to qualitative traits. In general, locating major genes is a relatively simple endeavor. Figure 2 provides an example of how to determine the distance between a molecular marker (RFLP) and a discrete phenotype (qualitative trait). In this exercise, a disease-resistant trait is used. Plants showing the presence and absence of a resistance trait are represented by R and r, respectively. For simplicity, 16 segregating backcross individuals are shown (Figure 2). The resistance locus is determined by identifying nearby molecular markers. M_1 represents marker allele 1, which segregates together with the disease-resistant (R) allele in parent 1 (P_1). M_2 represents marker allele 2, which seg-

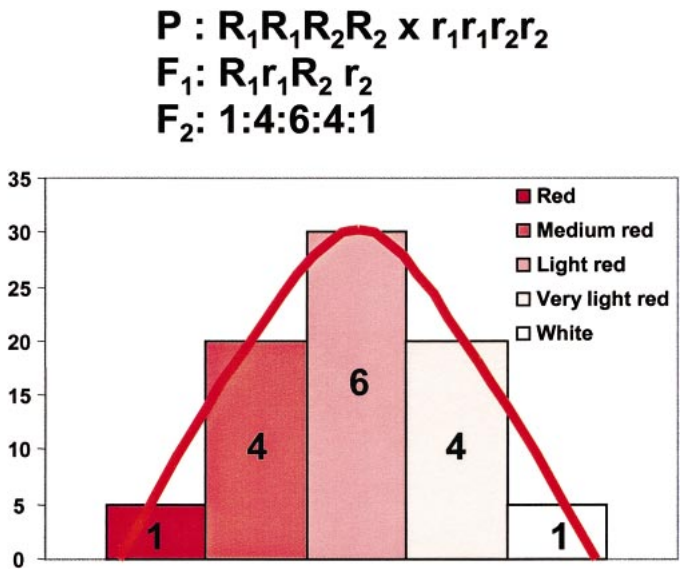

FIGURE 3. Quantitative vs. qualitative traits. Quantitative and qualitative traits are not fundamentally different. In this classical example, a homozygous dominant red strain was crossed with a homozygous recessive white strain ($R_1R_1R_2R_2 \times r_1r_1r_2r_2$), and a heterozygous ($R_1r_1R_2r_2$) light red F_1 is obtained. Selfing of the F_1 produces five different genotypes of F_2 with a ratio of 1:4:6:4:1. If the population of F_2 plants exhibit discrete phenotypic classes as red, medium red, light red, very light red, and white; this color trait is qualitative. However, if a rather symmetrical, bell-shaped histogram is obtained, the trait is quantitative.

regates together with the disease-susceptible allele (r) in parent 2 (P_2). A cross between two homozygous parents ($RM_1/RM_1 \times rM_2/rM_2$) produces a heterozygous (RM_1/rM_2) F_1 generation. A single crossover occurs during meiosis, which gives rise to four different gamete types (RM_1, RM_2, rM_1, and rM_2). The genotype of these gametes can be identified by a testcross between the F_1 and the disease-susceptible parent, P_2. Four classes of testcross progeny are obtained (RM_1/rM_2, RM_2/rM_2, rM_1/rM_2, and rM_2/rM_2), in which RM_2/rM_2 and rM_1/rM_2 are the recombinants. The genomic DNA of 16 backcross progeny are then digested with a restriction enzyme, run on an agarose gel, blotted onto a membrane, and hybridized with a probe that will detect the single fragment with an identical (or nearly identical) sequence. As shown in Figure 2B, two individual progeny (#2 and #10) have recombined the original linkage relationships of the parent. These two individuals are called crossovers, and the genetic distance (in centimorgans) between the locus R and M is determined by the percentage of crossover.

Qualitative vs. Quantitative Trait

A quantitative trait is much more complicated than a qualitative trait; the basic nature of quantitative traits is that polygenes and environmental factors contribute to phenotypic variability. However, qualitative and quantitative traits are not fundamentally different. The following classic example helps explain the difference (Figure 3). In 1910, Swedish geneticist Nilsson-Ehle (cited in Stansfield 1991) crossed two wheat (*Triticum aestivum* L.) strains, a red and a white seed, and produced light red plants in the F_1 generation. When F_1 plants were self-fertilized, among all F_2 plants, 1/16 was red and 1/16 was white. Thus, he postulated that there were two genes, each with a pair of alleles contributing to a cumulative effect. In this example, the use of capital and lowercase letters does not imply dominant and recessive allelic interactions, but rather additive gene action in which each dominant gene (R_1 or R_2) makes an equal contribution to redness, and each recessive gene (r_1 or r_2) contributes nothing. After crossing a homozygous dominant line with a homozygous recessive line ($R_1R_1R_2R_2 \times r_1r_1r_2r_2$), a dihybrid heterozygous F_1 ($R_1r_1R_2r_2$) is obtained. Selfing of the F_1 produces five different genotypes of F_2 with a ratio of 1:4:6:4:1. If the population of F_2 plants exhibit a discontinuous variation, we can discretely separate phenotypic classes as red, medium red, light red, very light red, and white (Figure 3). This color trait is qualitative. In contrast, if a rather symmetrical, bell-shaped histogram (Figure 3) is obtained, the trait is quantitative. Because this type of trait carries both qualitative and quantitative qualifications, it is also called a quasi-quantitative trait (Stansfield 1991).

QTL Mapping

Quantitative trait loci (QTL) cannot be mapped the same way as qualitative traits because the individual loci cannot be identified. The basis of all QTL mapping is to associate genetically determined phenotypes with molecular markers. Thus, rigorous statistical procedures are required and a detailed linkage map is a prerequisite for identifying QTL (Kearsey and Pooni 1996). Until the development of molecular markers, it was not practicable to locate quantitative traits using major gene mutants for a number of reasons. First, there are not many major gene mutants. Second, a major gene mutant is likely to have pleiotropic effects (a single mutation produces multiple phenotypic effects) on the quantitative trait and affects the fitness of the individuals that carry the mutation. Moreover, major gene mutants cause distortion of segregation in the marker locus, and thus affect the measurement of recombination. However, the identification of QTL has become possible with the advent of molecular markers and the development of genetic maps. There are many single locus polymorphisms at the DNA level. These markers segregate as single genes, and because they are naturally occurring, they are unlikely to affect fitness and thus will not distort segregation ratios. These loci can be identified with restriction enzymes in combination with marker probes (cDNA). Figure 4 illustrates the principle of QTL mapping based on the distance between a single QTL (responsible for seed germinability) and a nearby molecular marker.

Let us assume that a RFLP marker allele M_1 is linked to the germinability allele (G) for high-percentage germination in parent P_1, and M_2 is linked to the germinability allele (g) for low-percentage germination in parent P_2. A cross between these two homozygous parents ($GM_1/GM_1 \times gM_2/gM_2$) produces a heterozygous (GM_1/gM_2) F_1 generation. All the seeds of F_1 plants will be intermediate for percentage germination. The F_1 segregates and gives rise to four kinds of gametes (GM_1, GM_2, gM_1, and gM_2). Similar to the above example (Figure 2), the genotype of these gametes can be identified by carrying out a testcross between the F_1 and the parent P_2. Four classes of testcross progeny are obtained (GM_1/gM_2, GM_2/gM_2, gM_1/gM_2, and gM_2/gM_2), in which GM_2/gM_2 and gM_1/gM_2 are the recombinants (Figure 4A).

Figure 4B exhibits the germination distribution curves

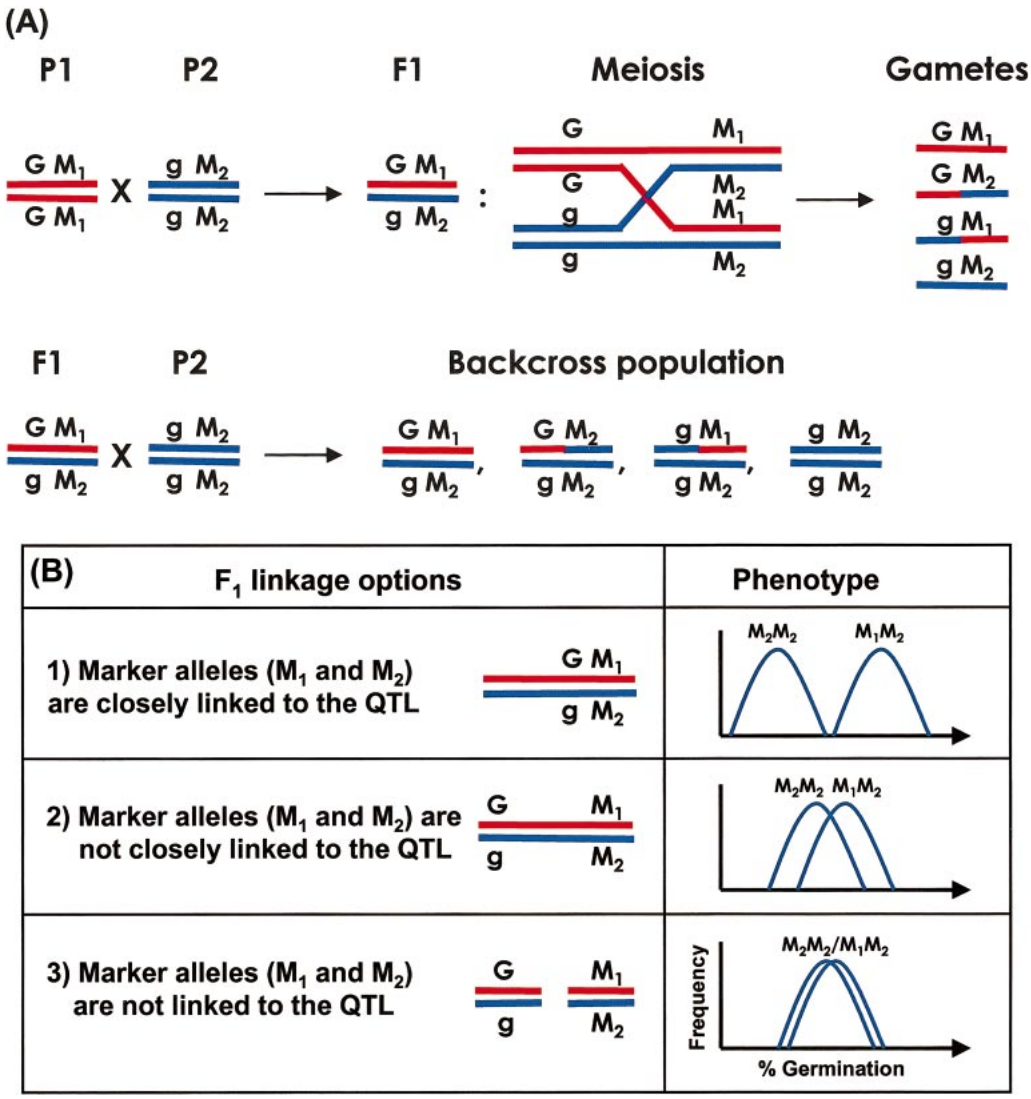

FIGURE 4. QTL mapping is to associate genetically determined phenotypes with DNA markers. (A) In this example, a single QTL responsible for seed germinability and a nearby molecular marker (RFLP) are used to illustrate the principle of QTL mapping. Marker allele M_1 is linked to the germinability allele (G) for high-percentage germination in parent P_1, and M_2 is linked to the germinability allele (g) for low-percentage germination in parent P_2. A cross between these two homozygous parents ($GM_1/GM_1 \times gM_2/gM_2$) produces a heterozygous (gM_1/gM_2) F_1 generation. The F_1 segregates and gives rise to four different gamete types (GM_1, GM_2, gM_1, and gM_2). The genotype of these gametes is identified by carrying out a testcross between F_1 and parent P_2. Four classes of testcross progeny are obtained (GM_1/gM_2, GM_2/gM_2, gM_1/gM_2, and gM_2/gM_2), in which the genotype of GM_2/gM_2 and gM_1/gM_2 are the recombinants. (B) *X*-axis (% germination) represents an increase in percentage germination. *Y*-axis (frequency) represents the number of plants carrying the marker allele M_1 or M_2. The key to mapping a QTL is to plot germination distribution curves separately for plants carrying the M_1 or M_2 marker allele. Three different types of germination distribution curves are shown based on the distance between marker and germinability alleles. (1) When the marker allele (M_1 and M_2) is tightly linked to the germinability allele (G and g), the majority of testcross progeny will be nonrecombinant types (GM_1/gM_2 and gM_2/gM_2), namely, seeds that show a high percentage of germination carry a M_1 allele and seeds that show a low percentage of germination carry only the M_2. Two distinct germination distribution curves are obtained. (2) When the marker allele is not closely linked to the germinability allele, a higher number of recombinant progeny will be obtained. Thus, two curves overlap each other. (3) When the marker allele is unlinked to the QTL, all four genotypes (GM_1/gM_2, GM_2/gM_2, gM_1/gM_2, and gM_2/gM_2) will occur at equal ratio in testcross progeny because of independent segregation. A single-germination distribution curve is obtained.

based on the distance between the marker and the germinability alleles. The *X*-axis represents an increase in percentage germination, and the *Y*-axis represents the number of plants carrying the marker allele M_1 or M_2. The key to mapping a QTL is to plot germination distribution curves separately for plants carrying the M_1 or M_2 marker allele (Figure 4B). For example, if the marker allele (M_1 and M_2) is tightly linked to the germinability allele (G and g), the majority of testcross progeny will be nonrecombinant types (GM_1/gM_2 and gM_2/gM_2). They will include seed that shows a high percentage of germination and carries the M_1 allele, and seed that shows a low percentage of germination and carries only the M_2. We will obtain two distinct germination distribution curves (Figure 4B-1). On the other hand, if the marker allele is distantly linked to the germinability allele, a higher number of recombinant progeny will be obtained, and these two curves will show some degree of overlap (Figure 4B-2). Furthermore, if the marker allele is unlinked to the QTL, all four genotypes (GM_1/gM_2, GM_2/gM_2, gM_1/gM_2, and gM_2/gM_2) will occur at an equal ratio in testcross progeny because of independent segregation. We will have a single-germination distribution curve (Figure 4B-

3). Thus, depending upon the distance between the marker and the germinability locus, the germination distribution curve is different and, the more markers we test, the better the chances of finding those closely linked to QTL (for more information, see Jones et al. 1997; Kearsey and Pooni 1996). The distance between the QTL and molecular markers can be estimated based on linear regression (Edwards et al. 1987), which examines the relationship between the performance for the quantitative trait and the genotypes at the marker locus. If there is a statistically significant association between the trait performance and the marker locus gene types, it is inferred that a QTL is located near the marker locus.

In reality, usually several QTL may account for germinability variation, and each QTL shares a small percentage of the total phenotypic variation for the trait. Two major problems are typically associated with QTL mapping. First, many individuals that are identical at a particular locus exhibit different phenotypes because of differences in environmental conditions or in genetic backgrounds. These phenomena are called penetrance. For a single-gene trait, biological and environmental limitation accounts for penetrance, but in a multigene trait, the genetic context also determines penetrance besides biological and environmental factors. Another problem associated with basic QTL analysis is that it cannot distinguish between tight linkage to a QTL with a small effect and loose linkage to a QTL with a large effect, resulting in great discrepancy between estimated and physical distance in a QTL map (Lander and Botstein 1989). The QTL interval mapping (by maximum likelihood) overcomes some of this problem. The principle of interval mapping is to test a model for the presence of a QTL at numerous positions between two known marker loci (Lander and Botstein 1989; Lincoln et al. 1993; Nelson 1997).

Mapping Population

Two types of mapping populations, F_2 and recombinant inbred (RI), are widely used for QTL mapping. An F_2 population is derived by selfing or intermating among F_1 individuals. The major advantage of an F_2 population is that the effects of additive and dominance gene actions at specific loci can be measured. RI lines are developed by repeated selfing and selection from individual plants of an F_2 population (Paterson 1996a). As most individuals, if not all, are homozygous in a RI population, all the seed from an individual plant can be bulked with near-perfect fidelity. This makes the RI a popular source for QTL mapping. The disadvantage of RI lines is that only additive gene action can be measured. Another mapping population, near-isogenic lines (NILs), is used to develop fine mapping of QTL. NILs are developed by crossing a donor line carrying a specific trait of interest to a recurrent line with desirable qualities. By repeatedly backcrossing a line carrying the dominant allele of the target gene to the recurrent parent, the donor genome is progressively eliminated, except for a small chromosomal segment containing the selected dominant allele. Thus, after several generations of backcrossing, the progeny becomes nearly isogenic with the recurrent parent (for more information, see Paterson 1996b; Tanksley et al. 1995).

NILs are good for cloning quantitative trait genes because we can specifically select a single QTL responsible for a major phenotypic effect. This approach is not only capable of developing a fine map near the target QTL, but also diminishes polygene-generated penetrance. Moreover, unlike other mapping populations, NILs exclude the phenotypic variability caused by genetic background. Recently, the first QTL (*fw2.2*) responsible for the major effect of tomato (*Lycopersicon esculentum* L.) fruit size (30%) was isolated by Alpert and Tanksley (1996) using the F_2 NIL mapping population, and the cDNA clone (*ORFX*) responsible for this trait was isolated later by Frary et al. (2000). Genetic complementation analysis showed that *ORFX* is expressed early in floral development and regulates carpel cell number (Frary et al. 2000). This result provides a paradigm of how a plant QTL can be cloned using NILs. Nevertheless, not every species is suitable to develop a NIL mapping population. Trees, for example, are not suitable because of their long generation interval. In this circumstance, an alternative candidate gene strategy may be applied (Rothschild and Soller 1997). This approach focuses on polymorphisms in, or close to, genes whose functions are directly related to the trait of interest. If an association between a candidate gene polymorphism and the trait of interest is established, it is assumed that the gene is involved in the regulation of that trait. Recently, Frewen et al. (2000) have mapped five candidate genes putatively involved in the regulation of bud dormancy and established that two genes (*PHYB2* and *ABI1B*) are coincident with QTL, affecting bud set and bud flush. However, because this approach is suggestive, the most challenging task is to verify if the candidate genes are true QTL.

Traditional Approach for Function Determination

Once a handful of cDNA clones is isolated, it is important to determine what functional roles the corresponding genes play in cellular activities. Depending upon the techniques used for gene identification and cloning, the methods for the determination of function differ. Traditionally, two approaches are used: forward genetics and reverse genetics. In the last 5 yr, a new approach that uses modified viral vectors has been developed.

Forward Genetics

Forward genetics starts out with a mutant plant and investigates which gene(s) gives rise to the altered phenotype. Thus, the locus responsible for a specific trait has been defined at the molecular level before initiating the cloning process. Most genes that are obtained via mutation-based approaches (QTL analysis, T-DNA knockouts, etc.) belong to this category. The function of corresponding wild-type cDNA clones are known. Thus, there is no need to further determine the function of these clones, except to confirm that they still carry out the expected activities by functional complementation. The most common way of performing functional complementation is to transform the wild-type gene or cDNA into the mutant plant, and check if the transgenic mutant plant restores its wild-type phenotype.

Reverse Genetics

Reverse genetics alters the gene sequence or expression pattern in vivo and then determines its function by evalu-

ating the change in the phenotype. If cDNA clones are obtained via expression-based technology, such as differential RNA display, differential screening of a cDNA library, or microarray, the functions of the corresponding genes are generally unknown. Thus, a reverse genetics approach is needed for function determination unless the sequence of the gene has a good match with the sequence of other genes whose function has been identified or the deduced amino acid sequence provides certain clues on a possible mechanistic role. Two methods are widely adopted by plant molecular biologists. One method is to generate a transgenic plant that down-regulates the expression of the endogenous target gene. A chimeric gene that expresses an antisense or double-stranded RNA in transgenic plants is constructed and transformed into the plant genome. The function of the gene can then be evaluated based on loss-of-function phenotype through suppressing the expression of the endogenous target gene if there is no functional redundancy within gene families. The other method is to generate a transgenic plant in strategies to overexpress a gene product. A chimeric gene is constructed by fusing a strong promoter with the gene of interest in the sense orientation and is transformed into the plant genome. Two types of phenotypic alterations can be anticipated: (1) the transgene mRNA is overexpressed, and thus gene function can be identified by the gene product's strong physiological role; and (2) both the transgene and endogenous target gene are suppressed (cosuppression) and the plants exhibit loss-of-function phenotype.

Viral Approach for Function Determination

Recent developments in virus–plant interactions have revealed that a viral vector modified with the host gene can be a powerful tool for investigating plant gene function. The merit of using viral vectors is speed, as we can inoculate viral constructs in plants directly and examine the altered phenotype shortly afterwards (Figure 5). The gene function is basically determined by a loss-of-function or an overexpressing phenotype (Baulcombe 1999; Fischer et al. 1999). The notion of forward or reverse genetics can also be applied to the viral approach. For example, if this technology is performed on a genome scale by generating a cDNA library, using one of the viral vectors and systematically screening for specific phenotype from large numbers of infected plants, it is referred to as forward genetics because the screening process is based fully on alteration of the phenotype; the information of gene sequence is not known beforehand. On the other hand, if the cDNA clones are isolated by microarray analysis, and the sequence of each gene is known, function identification using genetically modified viral vectors is then referred to as reverse genetics. Two distinct methods are currently used to determine gene function: virus-induced gene silencing (VIGS) (by loss-of-function phenotype) and virus-directed gene expression (by overexpressing phenotype) (Baulcombe 1999; Fischer et al. 1999).

Virus-induced Gene Silencing

In 1992, it was first shown that plants expressing untranslatable transcripts of tobacco (*Nicotiana tabacum* L.) etch virus coat–protein became resistant to viral infection from

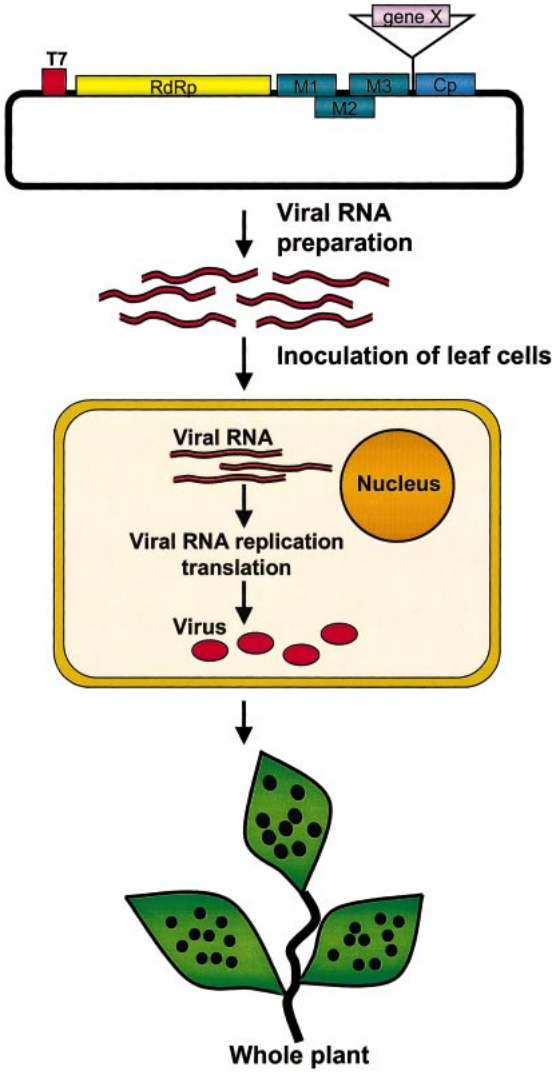

FIGURE 5. Viral approach for function determination. cDNA clone of potato virus X (Baulcombe et al. 1995) is used as an example. Features of this virus include RNA-dependent RNA polymerase gene (RdRp), genes encoding movement proteins (M1, M2, and M3), coat protein gene (CP), and segment of plant gene (gene X) to be studied. The 5′ end of the virus clone is flanked by a promoter from T7 bacteriophage (T7) which is part of the regular plasmid vector. Viral RNA is prepared in vitro and mechanically inoculated into the plant tissue. Upon inoculation, viral RNA is replicated in the cytoplasm, and viral proteins are synthesized. The virus moves systemically through the whole plant resulting in suppression of the homologous endogenous plant gene.

the same strain (Lindbo and Dougherty 1992). Different lines of evidence, however, showed that if plants are infected with genetically modified viral vectors carrying the fragment of a host gene, the expression of this gene is suppressed in the host plant (Baulcombe 1999). This phenomenon is called VIGS. Importantly, VIGS has been applied for rapid functional analysis of unknown genes (Burton et al. 2000; Kjemtrup et al. 1998; Kumagai et al. 1995; Ruiz et al. 1998). At least three different virus vectors have been used for functional studies in tobacco species. Genetically modified potato virus X has been used to silence an endogenous phytoene desaturase gene, a green fluorescent protein transgene (Ruiz et al. 1998), and an endogenous cellulose synthase gene (Burton et al. 2000). The geminivirus tomato golden mosaic virus has been used to suppress an endoge-

nous magnesium chelatase gene and a firefly luciferase transgene (Kjemtrup et al. 1998). Likewise, the tobacco mosaic virus carrying tomato phytoene desaturase repressed endogenous gene expression in *N. benthamiana* (Kumagai et al. 1995).

VIGS is likely an RNA-mediated defense in plants (Ratcliff et al. 1997). When a virus carries a homolog of a plant gene, the RNA-mediated defense mechanism targets both the viral RNA and the endogenous host RNA, resulting in gene silencing in the host plant. Recently, vast lines of evidence indicate that posttranscriptional gene silencing (PTGS) or double-stranded RNA interference (RNAi) plays the major role for VIGS (Anandalakshmi et al. 1998; Cogoni and Macino 2000; Kasschau and Carrington 1998; Waterhouse et al. 1999). Based on transgene-induced, virus-induced, or other transient assay-induced gene silencing, the occurrence of PTGS can be summarized in the following three interrelated models (Dalmay et al. 2000a; Matzke et al. 2000; Waterhouse et al. 1998, 1999; Yang et al. 2000) (Figure 6).

First, the duplex RNA model proposes that double-stranded RNA induces mRNA degradation rapidly (Waterhouse et al. 1998; Yang et al. 2000). The degradation product could be short RNA species of 21 to 23 nts (Dalmay et al. 2000a; Hamilton and Baulcombe 1999; Hutvágner et al. 2000; Yang et al. 2000). It is postulated that these short RNA species are components of the systemic signal and specificity determinants of PTGS (Dalmay et al. 2000a; Hamilton and Baulcombe 1999). Second, the threshold model proposes that plant cells have a surveillance system that induces transgene mRNA degradation when it detects an above-threshold concentration (Goodwin et al. 1996; Waterhouse et al. 1999). Here an RNA-dependent RNA polymerase may play a role by converting the single-stranded RNA into a double-stranded RNA (Dalmay et al. 2000b). The double-stranded form may be further processed into short 21 to 23 nt segments. Finally, methylation of the transgene coding sequence is often associated with posttranscriptional gene silencing. The mechanism is not fully understood. However, fragments of degraded RNA as an end product of PTGS may reenter the nucleus and facilitate homologous DNA methylation (Matzke et al. 2000). Methylation of the transgene may be responsible for the maintenance of PTGS (Dalmay et al. 2000a).

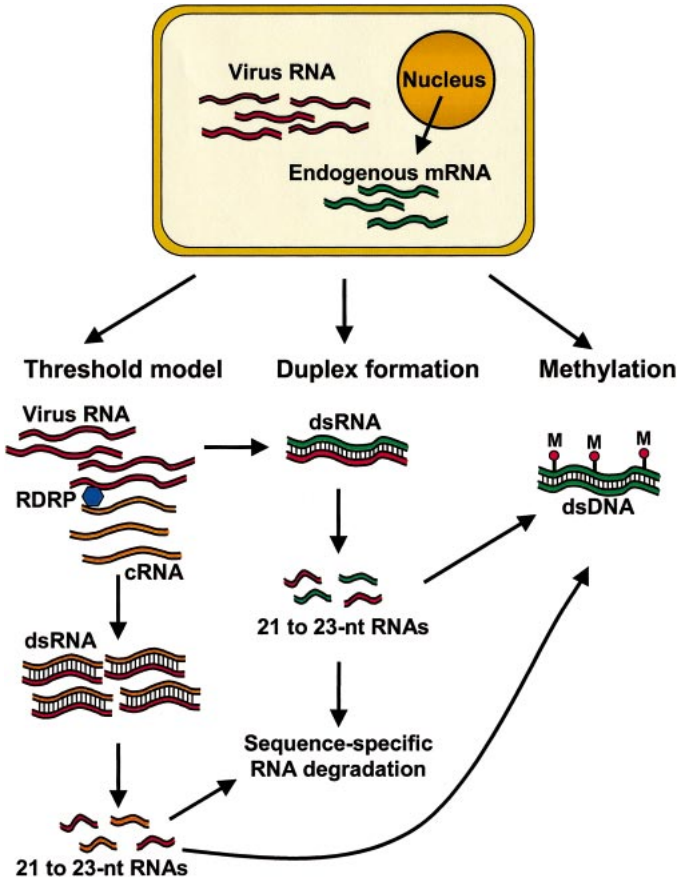

FIGURE 6. Three interacting models may explain the initiation and maintenance of virus-induced gene silencing. The duplex RNA model proposes that double-stranded RNA induces mRNA degradation rapidly. The degradation product could be short RNA species of 21 to 23 nts and components of the systemic signal and specificity determinants of posttranscriptional gene silencing (PTGS). The threshold model proposes that plant cells have a surveillance system that induces transgene mRNA degradation when it detects an above-threshold concentration of mRNA. An RNA-dependent RNA polymerase may convert the single-stranded RNA into a double-stranded RNA, which may be further processed into short 21 to 23 nt segments. Methylation of transgene coding sequence is often associated with posttranscriptional gene silencing. The mechanism is not fully understood. However, fragments of degraded RNA as an end product of PTGS may reenter the nucleus and facilitate homologous DNA methylation. Methylation of the transgene may be responsible for the maintenance of PTGS.

Virus-directed Gene Expression

Genetically modified viral vectors can also be used to investigate plant genes by transient expression of recombinant proteins. For gene expression studies, the constructs need to be designed to produce sense RNA. Current methods of inserting foreign genes into plant viral genomes include (1) gene replacement, where a nonessential viral gene such as coat protein is replaced by the gene of interest; (2) gene insertion, where the gene of interest is placed under the control of a strong subgenomic promoter; and (3) gene fusion, where the gene of interest is fused with a viral gene. The choice of methods depends on host–virus combinations, the target gene itself, and experimental objectives (Fischer et al. 1999). Using such recombinant viral vectors, Verch et al. (1998) has successfully expressed a monoclonal antibody in *N. benthamiana*. Currently, virus-directed gene expression in plants has attracted commercial interest. For example, Large Scale Biology (Vacaville, CA) uses modified tobacco mosaic viral vectors to screen several hundred genes per day for function (www.lsbc.com).

Conclusions

It is obvious that the goal of genomic research is to improve our lives by finding new methods for disease control, improving yield and quality of crops, enriching specific ingredients in our food, and so on. Likewise, investigating the molecular nature and cellular basis of signals that carry out the processes of regulating dormancy could help us to control the dormancy at will. Our lab studies leafy spurge, a perennial weed that causes persistent problems in the United States and Canada. Underground adventitious bud dormancy is the fundamental reason that leafy spurge escapes current control measures. Thus, new knowledge is needed to develop alternative strategies to improve the effectiveness of

existing control measures or to develop novel control strategies. Identification of various important gene components involved in root bud dormancy and growth is a way to approach a more efficient and effective manipulation of leafy spurge. For example, this information may allow scientists to develop new chemicals that can impair specific cellular processes or block critical pathways (Hess et al. 2001). In fact, most existing herbicides do target important cellular activities such as ACC-ase, the ALS enzyme, growth regulators, photosystem II, lipid synthesis, glutamine synthetase, PPO, and so forth.

Genetically modified viral vectors can be an alternative way of controlling weed growth based on the results of VIGS in plants (Burton et al. 2000; Kjemtrup et al. 1998; Kumagai et al. 1995; Ruiz et al. 1998). As a viral vector carrying host gene fragments may prevent the expression of homologous, chromosomal genes of the host, thus introducing a vital gene into the host plant via virus could inhibit its ability to grow. Besides, viruses can propagate themselves and systemically spread from tissue to tissue, perhaps avoiding the need for repeated applications. Moreover, this approach is specific because only those plants carrying highly similar genes such as those of an engineered virus would be affected. We are currently testing the lethality of a number of genes using a tomato bushy stunt virus gene vector, pHST2 (kindly provided by Dr. H. B. Scholthof, Texas A&M University, Texas) in *N. benthamiana*. Consequently, we may be able to reduce soil and air contamination if DNA, instead of herbicides, is used for weed management.

Acknowledgments

I thank Dr. Linda Walling for helpful comments on this manuscript and Dr. Xingyou Gu for providing a Southern blot picture (Figure 2B).

Literature Cited

Aharoni, A., L.C.P. Keizer, H. J. Bouwmeester et al. 2000. Identification of the *SAAT* gene involved in strawberry flavor biogenesis by use of DNA microarrays. Plant Cell 12:647–661.

Alpert, K. B. and S. D. Tanksley. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing *fw2.2*: a major fruit weight quantitative trait locus in tomato. Proc. Natl. Acad. Sci. USA 93:15 503–15 507.

Anandalakshmi, R., G. J. Pruss, X. Ge, R. Marathe, T. H. Smith, and V. B. Vance. 1998. A viral suppressor of gene silencing in plants. Proc. Natl. Acad. Sci. USA 95:13 079–13 084.

Anderson, J. V., W. S. Chao, and D. P. Horvath. 2001. A current review on the regulation of dormancy in vegetative buds. Weed Sci. 49:581–589.

Ausubel, F. M. 2000. *Arabidopsis* genome. A milestone in plant biology. Plant Physiol. 124:1451–1454.

Baulcombe, D. C. 1999. Fast forward genetics based on virus-induced gene silencing. Curr. Opin. Plant Biol. 2:109–113.

Baulcombe, D. C., S. Chapman, and S. Santa Cruz. 1995. Jellyfish green fluorescent protein as a reporter for virus infections. Plant J. 7:1045–1053.

Brown, S. M. and S. Kresovich. 1996. Molecular characterization for plant genetic resources conservation. Pages 86–93 *in* A. H. Patterson, ed. Genome Mapping in Plants. Austin: Academic Press/R. G. Landes.

Buchanan, B. B., W. Gruissem, and R. L. Jones. 2000. Biochemistry & Molecular Biology of Plants. Rockville, MD: American Society of Plant Physiologists. pp. 260–310.

Burton, R. A., D. M. Gibeaut, A. Bacic, K. Findlay, K. Roberts, A. Hamilton, D. C. Baulcombe, and G. B. Fincher. 2000. Virus-induced silencing of a plant cellulose synthase gene. Plant Cell 12:691–705.

Cogoni, C. and G. Macino. 2000. Post-transcriptional gene silencing across kingdoms. Curr. Opin. Genet. Dev. 10:638–643.

Dalmay, T., A. Hamilton, E. Mueller, and D. C. Baulcombe. 2000a. *Potato virus X* amplicons in *Arabidopsis* mediate genetic and epigenetic gene silencing. Plant Cell 12:369–379.

Dalmay, T., A. Hamilton, S. Rudd, S. Angell, and D. C. Baulcombe. 2000b. An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell 101:543–553.

Duggan, D. J., M. Bittner, Y. Chen, P. Meltzer, and J. M. Trent. 1999. Expression profiling using cDNA microarrays. Nat. Genet. Suppl. 21: 10–14.

Dutt, M. J. and K. H. Lee. 2000. Proteomic analysis. Curr. Opin. Biotechnol. 11:176–179.

Edwards, M. D., C. W. Stuber, and J. F. Wendel. 1987. Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125.

Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA 95:14 863–14 868.

Epstein, C. B. and R. A. Butow. 2000. Microarray technology–enhanced versatility, persistent challenge. Curr. Opin. Biotechnol. 11:36–41.

Fischer, R., C. Vaquero-Martin, M. Sack, J. Drossard, N. Emans, and U. Commandeur. 1999. Toward molecular farming in the future: transient protein expression in plants. Biotechnol. Appl. Biochem. 30: 113–116.

Foder, S. P. A., J. L. Read, M. C. Pirrung, L. Stryer, A. T. Lu, and D. Solas. 1991. Light-directed, spatially addressable parallel chemical synthesis. Science 251:767–773.

Foley, M. E. 2001. Seed dormancy: an update on terminology, physiological genetics, and quantitative trait loci regulating germinability. Weed Sci. 49:305–317.

Frary, A., T. C. Nesbitt, A. Frary et al. 2000. *fw2.2:* a quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88.

Frewen, B. E., T. H. H. Chen, G. T. Howe, J. Davis, A. Rohde, W. Boerjan, and H. D. Bradshaw, Jr. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud fluch in Populus. Genetics 154:837–845.

Goodwin, J., K. Chapman, S. Swaney, T. D. Parks, E. A. Wernsman, and W. G. Dougherty. 1996. Genetic and biochemical dissection of transgenic RNA-mediated virus resistance. Plant Cell 8:95–105.

Gygi, S. P., Y. Rochon, B. R. Franza, and R. Aebersold. 1999. Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. 19: 1720–1730.

Hamilton, A. and D. C. Baulcombe. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286:950–952.

Hess, D., R. J. Anderson, and J. D. Reagan. 2001. High throughput synthesis and screening: the partner of genomics for discovery of new chemicals for agriculture. Weed Sci. 49:249–256.

Hutvágner, G., L. Mlynárová, and J.-P. Nap. 2000. Detailed characterization of the posttranscriptional gene-silencing-related small RNA in a GUS gene-silenced tobacco. RNA 6:1445–1454.

Jones, N., H. Ougham, and H. Thomas. 1997. Markers and mapping: we are all geneticists now. *In* H. Thomas and J. F. Farrar, eds. Putting Plant Physiology on the Map: Genetic Analysis of Developmental and Adaptive Traits. Cambridge: Cambridge University Press. [See also New Phytol. 137:165–177.]

Kasschau, K. D. and J. C. Carrington. 1998. A counter-defensive strategy of plant viruses: suppression of post-transcriptional gene silencing. Cell 95:461–470.

Kearsey, M. J. and H. S. Pooni. 1996. The Genetical Analysis of Quantitative Traits. London: Chapman and Hall. pp. 133–164.

Kjemtrup, S., K. S. Sampson, C. G. Peele, L. V. Nguyen, M. A. Conkling, W. F. Thompson, and D. Robertson. 1998. Gene silencing from plant DNA carried by a geminivirus. Plant J. 14:91–100.

Kumagai, M. H., J. Donson, G. Della-Cioppa, D. Harvey, K. Hanley, and L. K. Grill. 1995. Cytoplasmic inhibition of carotenoid biosynthesis with virus-derived RNA. Proc. Natl. Acad. Sci. USA 92:1679–1683.

Krysan, P. J., J. C. Young, and M. R. Sussman. 1999. T-DNA as an insertional mutagen in *Arabidopsis*. Plant Cell 11:2283–2290.

Lander, E. S. and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–189.

Lewin, B. 1997. Genes VI. Oxford: Oxford University Press. pp. 49–237.

Li, F., E. S. Barnathan, and K. Kariko. 1994. Rapid method for screening and cloning cDNAs generated in differential messenger RNA display

and application of northern blot for affinity capture of cDNAs. Nucleic Acids Res. 22:1764–1765.
Liang, P. and A. B. Pardee. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967–971.
Liang, P. and A. B. Pardee. 1995. Recent advances in differential display. Curr. Opin. Immunol. 7:247–280.
Lincoln, S., M. Daly, and E. Lander. 1993. Mapping Genes Controlling Quantitative Traits with MAPMAKER/QTL 1.1: A Tutorial and Reference Manual. Whitehead Institute Technical Report.
Lindbo, J. A. and W. G. Dougherty. 1992. Untranslatable transcripts of the tobacco etch virus coat protein gene sequence can interfere with tobacco etch virus replication in transgenic plants and protoplasts. Virology 189:725–733.
Matzke, M. A., M. F. Mette, and A. J. M. Matzke. 2000. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol. Biol. 43:401–415.
Nelson, J. C. 1997. QGENE: software for marker-based genomic analysis and breeding. Mol. Breed. 3:239–245.
Paterson, A. H. 1996a. Making genetic maps. Pages 23–39 *in* A. H. Patterson, ed. Genome Mapping in Plants. Austin: Academic Press/R. G. Landes.
Paterson, A. H. 1996b. Mapping genes responsible for differences in phenotype. Pages 41–54 *in* A. H. Patterson, ed. Genome Mapping in Plants. Austin: Academic Press/R. G. Landes.
Paterson, A. H., K. F. Schertz, Y.-R. Lin, S.-C. Liu, and Y.-L. Chang. 1995. The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. Proc. Natl. Acad. Sci. USA 92:6127–6131.
Ratcliff, F., B. D. Harrison, and D. C. Baulcombe. 1997. A similarity between viral defense and gene silencing in plants. Science 276:1558–1560.
Reymond, P., H. Weber, M. Damond, and E. E. Farmer. 2000. Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. Plant Cell 12:707–719.
Rothschild, M. F. and M. Soller. 1997. Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. Probe 8:13–20.
Ruiz, M. T., O. Voinnet, and D. C. Baulcombe. 1998. Initiation and maintenance of virus-induced gene silencing. Plant Cell 10:937–946.
Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press.
Schena, M., D. Shalon, R. W. Davis, and P. O. Brown. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470.
Springer, P. S. 2000. Gene traps: tools for plant development and genomics. Plant Cell 12:1007–1020.
Stansfield, W. D. 1991. Schaum's Outline of Theory and Problems of Genetics. 3rd ed. Schaum's Outline Series. New York: McGraw-Hill. pp. 209–212.
Stevenson, B. J. 1999. Available: LICR Office of Information Technology, http://epd.unil.ch/biocomputing/array/software/MicroArray_Software.htm.
Sun, Y., G. Hegamyer, and N. H. Colburn. 1994. Molecular cloning of 5 messenger RNAs differentially expressed in preneoplastic or neoplastic JB6 mouse epidermal cells—one which is homologous to human tissue inhibitor of metalloproteinase 3. Cancer Res. 54:1139–1144.
Tanksley, S. D., M. W. Ganal, and G. B. Martin. 1995. Chromosome landing: a paradigm for map-based gene cloning in plants with large genomes. Trends Genet. 11:63–68.
Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res. 17:6463–6471.
Tavazoie, S., J. D. Hugues, M. J. Campbell, R. J. Cho, and G. M. Church. 1999. Systematic determination of genetic network architecture. Nat. Genet. 22:281–285.
van Hal, N. L. W., O. Vorst, A. M. M. L. van Houwelingen, E. J. Kok, A. Peijnenburg, A. Aharoni, A. J. van Tunen, and J. Keijer. 2000. The application of microarrays in gene expression analysis. J. Biotechnol. 78:271–280.
Veculescu, V. E., L. Zhang, B. Vogelstein, and K. W. Kinzler. 1995. Serial analysis of gene expression. Science 270:484–487.
Verch, T., V. Yusibov, and H. Koprowski. 1998. Expression and assembly of a full-length monoclonal antibody in plants using a plant virus vector. J. Immunol. Methods 220:69–75.
Waterhouse, P. M., M. W. Graham, and M.-B. Wang. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. USA 95:13 959–13 964.
Waterhouse, P. M., N. A. Smith, and M.-B. Wang. 1999. Virus resistance and gene silencing: killing the messenger. Trends Plant Sci. 4:452–457.
Wen, X., S. Fuhrman, G. S. Michaels, D. B. Carr, S. Smith, J. L. Barker, and R. Somogyi. 1998. Large-scale temporal gene expression mapping of central nervous system development. Proc. Natl. Acad. Sci. USA 95:334–339.
Williams, J. G., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535.
Yang, D., H. Lu, and J. W. Erickson. 2000. Evidence that processed small dsRNSs may mediate sequence-specific mRNA degradation during RNAi in *Drosophila* embryos. Curr. Biol. 10:1191–1200.
Zabeau, M. and P. Vos, inventors. 1993. Selective restriction fragment amplification: a general method for DNA fingerprinting. EPO patent 0534858A1.

Received May 8, 2001, and approved August 15, 2001.